

150134

ACCESS DB #

FOR OFFICIAL USE ONLY

PLEASE PRINT CLEARLY

Scientific and Technical Information Center

Checked JFR
5-19-2006

SEARCH REQUEST FORM

Requester's Full Name: Jeffrey E. Russel Examiner #: 62785 Date: 4-7-2005
 Art Unit: 1654 Phone Number: 2-0969 Serial Number: 10/674,516
 Location (Bldg/Room#): Room 3019 (Mailbox #): 3018 Results Format Preferred (circle): PAPER DISK

To ensure an efficient and quality search, please attach a copy of the cover sheet, claims, and abstract or fill out the following:

Title of Invention: Novel Peptide With Osteogenic Activity
 Inventors (please provide full names): S. Dhanaraj, A. Gosiewska, A. Reznia, G. Heavner,
X. Lin, C. Yi

Earliest Priority Date: 9-10-2004

Search Topic:

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known.

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search SEQ ID NO: 1 (CGGGRWCG) in STN, in the U.S. patent appl. sequence database, and in Genbank/Uniprot/PIR.

Thank you
JFR

STAFF USE ONLY

Searcher: _____

Searcher Phone #: _____

Searcher Location: _____

Date Searcher Picked Up: _____

Date Completed: _____

Searcher Prep & Review Time: _____

Online Time: _____

Type of Search

____ NA Sequence (#)

____ AA Sequence (#)

____ Structure (#)

____ Bibliographic

____ Litigation

____ Fulltext

____ Other

Vendors and cost where applicable

____ STN _____ Dialog

____ Questel/Orbit _____ Lexis/Nexis

____ Westlaw _____ WWW/Internet

____ In-house sequence systems

____ Commercial _____ Oligomer _____ Score/Length

____ Interference _____ SPDI _____ Encode/Transl

____ Other (specify)

=> d his full

(FILE 'HOME' ENTERED AT 16:45:36 ON 12 APR 2005)

FILE 'LREGISTRY' ENTERED AT 16:45:45 ON 12 APR 2005

L1 0 SEA ABB=ON CGGGRWCG/SQSP

FILE 'REGISTRY' ENTERED AT 16:46:05 ON 12 APR 2005

L2 0 SEA ABB=ON CGGGRWCG/SQSP

L3 44 SEA ABB=ON CGGGRWCG/SQSFP

SAVE TEMP L2 RUS649SEQ1/A

SAVE TEMP L3 RUS649SEQ2/A

FILE 'CAPLUS' ENTERED AT 16:46:40 ON 12 APR 2005

L4 33 SEA ABB=ON L3

FILE 'REGISTRY' ENTERED AT 16:47:15 ON 12 APR 2005

D QUE L3

D RN CN SQL KWIC NTE LC 1-44 L3

FILE 'CAPLUS, USPATFULL, USPAT2, TOXCENTER' ENTERED AT 17:17:26 ON 12 APR 2005

L5 40 SEA ABB=ON L3

L6 25 DUP REM L5 (15 DUPLICATES REMOVED)

ANSWERS '1-25' FROM FILE CAPLUS

D IBIB ED ABS HITRN 1-25

FILE 'HOME' ENTERED AT 17:18:01 ON 12 APR 2005

FILE HOME

FILE LREGISTRY

LREGISTRY IS A STATIC LEARNING FILE

FILE REGISTRY

Property values tagged with IC are from the ZIC/VINITI data file provided by InfoChem.

STRUCTURE FILE UPDATES: 11 APR 2005 HIGHEST RN 848290-51-7

DICTIONARY FILE UPDATES: 11 APR 2005 HIGHEST RN 848290-51-7

TSKA INFORMATION NOW CURRENT THROUGH JANUARY 18, 2005

Please note that search-term pricing does apply when conducting SmartSELECT searches.

*
* The CA roles and document type information have been removed from *
* the IDE default display format and the ED field has been added, *
* effective March 20, 2005. A new display format, IDERL, is now *
* available and contains the CA role and document type information. *
*

Crossover limits have been increased. See HELP CROSSOVER for details.

Checked
JRL
5-19-2006

Experimental and calculated property data are now available. For more information enter HELP PROP at an arrow prompt in the file or refer to the file summary sheet on the web at:
<http://www.cas.org/ONLINE/DBSS/registryss.html>

FILE CAPLUS

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

FILE COVERS 1907 - 12 Apr 2005 VOL 142 ISS 16
FILE LAST UPDATED: 11 Apr 2005 (20050411/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

FILE USPATFULL

FILE COVERS 1971 TO PATENT PUBLICATION DATE: 7 Apr 2005 (20050407/PD)
FILE LAST UPDATED: 7 Apr 2005 (20050407/ED)
HIGHEST GRANTED PATENT NUMBER: US6877166
HIGHEST APPLICATION PUBLICATION NUMBER: US2005076416
CA INDEXING IS CURRENT THROUGH 7 Apr 2005 (20050407/UPCA)
ISSUE CLASS FIELDS (/INCL) CURRENT THROUGH: 7 Apr 2005 (20050407/PD)
REVISED CLASS FIELDS (/NCL) LAST RELOADED: Feb 2005
USPTO MANUAL OF CLASSIFICATIONS THESAURUS ISSUE DATE: Feb 2005

```
>>> USPAT2 is now available. USPATFULL contains full text of the      <<<
>>> original, i.e., the earliest published granted patents or        <<<
>>> applications. USPAT2 contains full text of the latest US        <<<
>>> publications, starting in 2001, for the inventions covered in    <<<
>>> USPATFULL. A USPATFULL record contains not only the original    <<<
>>> published document but also a list of any subsequent            <<<
>>> publications. The publication number, patent kind code, and     <<<
>>> publication date for all the US publications for an invention   <<<
>>> are displayed in the PI (Patent Information) field of USPATFULL <<<
>>> records and may be searched in standard search fields, e.g., /PN, <<<
>>> /PK, etc.                                                         <<<

>>> USPATFULL and USPAT2 can be accessed and searched together      <<<
>>> through the new cluster USPATALL. Type FILE USPATALL to         <<<
>>> enter this cluster.                                              <<<
>>>                                                                    <<<
>>> Use USPATALL when searching terms such as patent assignees,     <<<
>>> classifications, or claims, that may potentially change from   <<<
>>> the earliest to the latest publication.                          <<<
```

This file contains CAS Registry Numbers for easy and accurate substance identification.

FILE USPAT2

FILE COVERS 2001 TO PUBLICATION DATE: 7 Apr 2005 (20050407/PD)

FILE LAST UPDATED: 7 Apr 2005 (20050407/ED)
HIGHEST GRANTED PATENT NUMBER: US2004265802
HIGHEST APPLICATION PUBLICATION NUMBER: US2005075819
CA INDEXING IS CURRENT THROUGH 5 Apr 2005 (20050405/UPCA)
ISSUE CLASS FIELDS (/INCL) CURRENT THROUGH: 7 Apr 2005 (20050407/PD)
REVISED CLASS FIELDS (/NCL) LAST RELOADED: Feb 2005
USPTO MANUAL OF CLASSIFICATIONS THESAURUS ISSUE DATE: Feb 2005

USPAT2 is a companion file to USPATFULL. USPAT2 contains full text of the latest US publications, starting in 2001, for the inventions covered in USPATFULL. USPATFULL contains full text of the original published US patents from 1971 to date and the original applications from 2001. In addition, a USPATFULL record for an invention contains a complete list of publications that may be searched in standard search fields, e.g., /PN, /PK, etc.

USPATFULL and USPAT2 can be accessed and searched together through the new cluster USPATALL. Type FILE USPATALL to enter this cluster.

Use USPATALL when searching terms such as patent assignees, classifications, or claims, that may potentially change from the earliest to the latest publication.

FILE TOXCENTER

FILE COVERS 1907 TO 5 Apr 2005 (20050405/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

TOXCENTER has been enhanced with new files segments and search fields. See HELP CONTENT for more information.

TOXCENTER thesauri in the /CN, /CT, and /MN fields incorporate the MeSH 2005 vocabulary. See <http://www.nlm.nih.gov/mesh/> and http://www.nlm.nih.gov/pubs/techbull/nd04/nd04_mesh.html for a description of changes.

=>

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 18:29:28 ; Search time 69 Seconds
 (without alignments)
 44.842 Million cell updates/sec

Title: PCT-US04-29649A-1
 Perfect score: 58
 Sequence: 1 CGGGRWCG 8

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
 1: geneseqp1980s:*
 2: geneseqp1990s:*
 3: geneseqp2000s:*
 4: geneseqp2001s:*
 5: geneseqp2002s:*
 6: geneseqp2003as:*
 7: geneseqp2003bs:*
 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		Length	DB	ID	Description
		Match					
1	49	84.5		85	4	AAU42392	Aau42392 Propionib
2	49	84.5		85	6	ABM38911	Abm38911 Propionib
3	45	77.6		269	3	AAV71010	Aay71010 Chicken p
4	45	77.6		432	1	AAP92149	Aap92149 Large env
5	45	77.6		432	1	AAP92148	Aap92148 Large env
6	45	77.6		662	2	AAW68401	Aaw68401 FeLV-B en
7	44	75.9		108	4	AAU65014	Aau65014 Propionib
8	44	75.9		108	6	ABM61533	Abm61533 Propionib
9	44	75.9		884	4	ABG09159	Abg09159 Novel hum

10	43	74.1	203	4	ABG11677	Abg11677	Novel hum
11	43	74.1	207	4	ABG06236	Abg06236	Novel hum
12	43	74.1	262	4	ABB65083	Abb65083	Drosophil
13	43	74.1	296	7	ABO68792	Abo68792	Pseudomon
14	43	74.1	500	7	ADD47142	Add47142	Rat Prote
15	43	74.1	500	7	ADD47146	Add47146	Rat Prote
16	43	74.1	501	7	ADF30455	Adf30455	Rat angio
17	43	74.1	559	8	ABM84024	Abm84024	Human dia
18	43	74.1	606	5	ABP52121	Abp52121	Homo sapi
19	43	74.1	620	7	ADE10036	Adel0036	Novel pro
20	43	74.1	2338	8	ABO58348	Abo58348	Human gen
21	43	74.1	2367	4	AAU31850	Aau31850	Novel hum
22	43	74.1	19938	6	ABP76678	Abp76678	Streptomy
23	42	72.4	151	4	AAU45804	Aau45804	Propionib
24	42	72.4	151	6	ABM42323	Abm42323	Propionib
25	42	72.4	269	7	ABO80763	Abo80763	Pseudomon
26	41	70.7	32	8	ADK01482	Adk01482	Hepatitis
27	41	70.7	109	6	ABU70643	Abu70643	Human adi
28	41	70.7	180	4	AAB95532	Aab95532	Human pro
29	41	70.7	727	2	AAW78435	Aaw78435	Human ADA
30	41	70.7	896	3	AAB21265	Aab21265	Mouse met
31	41	70.7	949	7	ADG72483	Adg72483	Human agg
32	41	70.7	950	2	AAU49501	Aay49501	Human MET
33	41	70.7	950	3	AAU53899	Aay53899	Amino aci
34	41	70.7	950	4	AAB73549	Aab73549	Human ADA
35	41	70.7	950	4	AAB50002	Aab50002	Human MET
36	41	70.7	950	8	ADO20218	Ado20218	Human PRO
37	41	70.7	950	8	ADQ39942	Adq39942	Human myo
38	41	70.7	967	2	AAW80285	Aaw80285	Human int
39	41	70.7	967	2	AAW78189	Aaw78189	Human sec
40	41	70.7	967	2	AAU04142	Aay04142	Human Tan
41	41	70.7	967	6	ADA57139	Ada57139	Human sec
42	41	70.7	967	6	ADA41003	Ada41003	Human sec
43	41	70.7	967	7	ADB91631	Adb91631	Human sec
44	41	70.7	967	7	ADC74267	Adc74267	Human sec
45	41	70.7	967	7	ADD37948	Add37948	Human sec

ALIGNMENTS

RESULT 1

AAU42392

ID AAU42392 standard; protein; 85 AA.

XX

AC AAU42392;

XX

DT 27-FEB-2002 (first entry)

XX

DE Propionibacterium acnes immunogenic protein #3288.

XX

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

KW dermatological; osteopathic; neuroprotectant.

XX

OS Propionibacterium acnes.

XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59517.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Example 1; SEQ ID NO 3587; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 85 AA;

Query Match 84.5%; Score 49; DB 4; Length 85;
 Best Local Similarity 87.5%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGRWCG 8
 || |||||
 Db 15 CGYGRWCG 22

Search completed: April 12, 2005, 18:57:00
Job time : 72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 18:54:49 ; Search time 41 Seconds
(without alignments)
14.566 Million cell updates/sec

Title: PCT-US04-29649A-1
Perfect score: 58
Sequence: 1 CGGGRWCG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	43	74.1	296	4	US-09-252-991A-17538
2	43	74.1	447	4	US-09-949-016-6956
3	43	74.1	563	4	US-09-949-016-6955
4	43	74.1	620	4	US-09-949-016-7495
5	43	74.1	620	4	US-09-949-016-7496
6	43	74.1	620	4	US-09-949-016-7497
7	43	74.1	620	4	US-09-949-016-7498
8	42	72.4	269	4	US-09-252-991A-29509

9	41	70.7	551	4	US-09-130-491-16	Sequence 16, Appl
10	41	70.7	608	4	US-09-130-491-13	Sequence 13, Appl
11	41	70.7	727	4	US-09-445-023A-1	Sequence 1, Appli
12	41	70.7	727	4	US-09-445-023A-12	Sequence 12, Appl
13	41	70.7	949	4	US-09-568-559-2	Sequence 2, Appli
14	41	70.7	950	4	US-09-321-987B-4	Sequence 4, Appli
15	41	70.7	967	4	US-09-130-491-2	Sequence 2, Appli
16	40.5	69.8	75	4	US-09-621-976-5341	Sequence 5341, Ap
17	39	67.2	70	4	US-09-248-796A-21926	Sequence 21926, A
18	39	67.2	147	3	US-09-134-001C-2854	Sequence 2854, Ap
19	39	67.2	439	4	US-09-902-540-13020	Sequence 13020, A
20	39	67.2	3854	4	US-09-949-016-7876	Sequence 7876, Ap
21	38	65.5	120	4	US-09-252-991A-24786	Sequence 24786, A
22	38	65.5	122	1	US-07-734-534A-1	Sequence 1, Appli
23	38	65.5	208	4	US-09-252-991A-25609	Sequence 25609, A
24	38	65.5	352	4	US-09-902-540-9796	Sequence 9796, Ap
25	38	65.5	416	4	US-09-540-236-2215	Sequence 2215, Ap
26	38	65.5	635	4	US-09-270-767-44609	Sequence 44609, A
27	38	65.5	2088	4	US-09-548-372D-13	Sequence 13, Appl
28	38	65.5	2088	4	US-09-548-367D-13	Sequence 13, Appl
29	38	65.5	2088	4	US-09-551-853D-13	Sequence 13, Appl
30	38	65.5	2088	4	US-09-548-376D-13	Sequence 13, Appl
31	38	65.5	2088	4	US-09-548-373D-13	Sequence 13, Appl
32	38	65.5	2088	4	US-09-548-366F-13	Sequence 13, Appl
33	38	65.5	2088	4	US-09-548-368D-13	Sequence 13, Appl
34	38	65.5	2508	4	US-09-627-650B-7	Sequence 7, Appli
35	38	65.5	2508	4	US-09-436-063C-7	Sequence 7, Appli
36	38	65.5	2544	4	US-09-627-650B-3	Sequence 3, Appli
37	38	65.5	2544	4	US-09-436-063C-3	Sequence 3, Appli
38	38	65.5	2601	4	US-09-627-650B-9	Sequence 9, Appli
39	38	65.5	2601	4	US-09-436-063C-9	Sequence 9, Appli
40	38	65.5	2732	4	US-09-086-436-30	Sequence 30, Appl
41	37.5	64.7	113	4	US-09-252-991A-29076	Sequence 29076, A
42	37	63.8	9	1	US-08-462-661A-110	Sequence 110, App
43	37	63.8	9	6	5318899-76	Patent No. 5318899
44	37	63.8	9	6	5318899-76	Patent No. 5318899
45	37	63.8	18	2	US-08-752-852A-151	Sequence 151, App

ALIGNMENTS

RESULT 1

US-09-252-991A-17538

; Sequence 17538, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17538
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17538

Query Match 74.1%; Score 43; DB 4; Length 296;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGRWC 7
|||
Db 70 GGGRWC 75

Search completed: April 12, 2005, 18:59:09
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 18:55:54 ; Search time 133 Seconds
(without alignments)
19.970 Million cell updates/sec

Title: PCT-US04-29649A-1
Perfect score: 58
Sequence: 1 CGGGRWCG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 331997259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	%		Query			ID	Description
	Score	Match	Length	DB			
1	49	84.5	225	16	US-10-437-963-106003	Sequence 106003,	
2	48	82.8	308	16	US-10-437-963-156549	Sequence 156549,	
3	46	79.3	114	16	US-10-437-963-108422	Sequence 108422,	
4	45	77.6	83	16	US-10-437-963-161633	Sequence 161633,	
5	45	77.6	117	16	US-10-437-963-120602	Sequence 120602,	
6	45	77.6	269	16	US-10-705-791-9	Sequence 9, Appli	
7	45	77.6	416	16	US-10-437-963-164329	Sequence 164329,	
8	44	75.9	59	16	US-10-437-963-113259	Sequence 113259,	
9	43.5	75.0	169	15	US-10-425-114-42832	Sequence 42832, A	
10	43	74.1	83	16	US-10-437-963-112746	Sequence 112746,	
11	43	74.1	96	16	US-10-767-701-56720	Sequence 56720, A	
12	43	74.1	97	15	US-10-425-114-48971	Sequence 48971, A	
13	43	74.1	103	16	US-10-437-963-186614	Sequence 186614,	
14	43	74.1	136	16	US-10-437-963-203306	Sequence 203306,	
15	43	74.1	203	16	US-10-437-963-120411	Sequence 120411,	
16	43	74.1	305	16	US-10-437-963-122157	Sequence 122157,	
17	43	74.1	501	14	US-10-316-253-18	Sequence 18, Appl	
18	43	74.1	539	16	US-10-437-963-172380	Sequence 172380,	
19	43	74.1	623	16	US-10-437-963-148905	Sequence 148905,	
20	43	74.1	2338	14	US-10-029-386-31982	Sequence 31982, A	
21	43	74.1	19725	15	US-10-084-846A-4	Sequence 4, Appli	
22	42	72.4	201	16	US-10-767-701-57762	Sequence 57762, A	
23	42	72.4	229	16	US-10-437-963-160205	Sequence 160205,	
24	42	72.4	411	15	US-10-425-114-67026	Sequence 67026, A	
25	42	72.4	709	16	US-10-437-963-174410	Sequence 174410,	
26	41	70.7	71	15	US-10-424-599-205131	Sequence 205131,	
27	41	70.7	74	16	US-10-437-963-107271	Sequence 107271,	
28	41	70.7	134	15	US-10-424-599-266046	Sequence 266046,	
29	41	70.7	139	16	US-10-767-701-32287	Sequence 32287, A	
30	41	70.7	161	16	US-10-437-963-128469	Sequence 128469,	
31	41	70.7	163	16	US-10-437-963-198571	Sequence 198571,	
32	41	70.7	215	16	US-10-437-963-155473	Sequence 155473,	
33	41	70.7	304	16	US-10-437-963-175382	Sequence 175382,	
34	41	70.7	420	16	US-10-437-963-145558	Sequence 145558,	

35	41	70.7	518	9	US-09-803-589-10	Sequence 10, Appl
36	41	70.7	543	16	US-10-437-963-196097	Sequence 196097,
37	41	70.7	551	9	US-09-802-582-16	Sequence 16, Appl
38	41	70.7	551	13	US-10-105-929-16	Sequence 16, Appl
39	41	70.7	551	14	US-10-365-227-16	Sequence 16, Appl
40	41	70.7	600	16	US-10-437-963-193474	Sequence 193474,
41	41	70.7	608	9	US-09-803-589-2	Sequence 2, Appli
42	41	70.7	608	9	US-09-803-589-8	Sequence 8, Appli
43	41	70.7	608	13	US-10-105-929-13	Sequence 13, Appl
44	41	70.7	727	9	US-09-445-023A-1	Sequence 1, Appli
45	41	70.7	727	9	US-09-445-023A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-10-437-963-106003
; Sequence 106003, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules
Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 106003
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_10488C.1.pep
US-10-437-963-106003

Query Match 84.5%; Score 49; DB 16; Length 225;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGRWCG 8
|||||||
Db 22 GGGRWCG 28

Search completed: April 12, 2005, 19:01:28
Job time : 134 secs

OM protein - protein search, using sw model

Run on: April 12, 2005, 18:49:54 ; Search time 16 Seconds
 (without alignments)
 48.108 Million cell updates/sec

Title: PCT-US04-29649A-1
 Perfect score: 58
 Sequence: 1 CGGGRWCG 8

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : PIR_79:*
 1: pir1:*
 2: pir2:*
 3: pir3:*
 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%				
No.	Score	Query Match	Length	DB	ID	Description
1	49	84.5	185	2	S12205	hypothetical prote
2	45	77.6	273	2	B31479	env polyprotein pr
3	45	77.6	347	2	S40105	gene AGP2beta-2 pr
4	45	77.6	534	1	VCVWSF	env polyprotein -
5	45	77.6	662	1	VCMVLB	env polyprotein -
6	45	77.6	662	1	VCVWGF	env polyprotein -
7	45	77.6	662	2	A25982	env polyprotein -
8	45	77.6	668	1	VCMVFP	env polyprotein pr
9	43	74.1	118	2	S52855	hypothetical prote
10	43	74.1	266	2	E64471	hypothetical prote
11	43	74.1	501	1	A32616	aldehyde dehydroge
12	43	74.1	501	1	JQ1004	aldehyde dehydroge
13	43	74.1	501	2	JC4524	aldehyde dehydroge
14	43	74.1	501	2	JC5553	aldehyde dehydroge

15	43	74.1	606	2	JC5604	ABC-transporting p
16	42	72.4	535	2	S58224	oestrogen receptor
17	41	70.7	550	2	T47158	hypothetical prote
18	41	70.7	951	2	T00017	gene ADAMTS-1 prot
19	40	69.0	241	2	C75329	conserved hypothet
20	40	69.0	461	2	T06268	probable beta-1,3-
21	39.5	68.1	368	2	G96668	protein F1N19.7 [i
22	39	67.2	54	2	A26463	relaxin - spiny do
23	39	67.2	295	2	JC5559	lectin-B - Virgini
24	38	65.5	138	1	I51381	phospholipase A2 (
25	38	65.5	138	1	PSRSB2	phospholipase A2 (
26	38	65.5	138	1	PSRSBT	phospholipase A2 (
27	38	65.5	264	2	B89005	protein T24A6.18 [
28	38	65.5	385	2	T20410	hypothetical prote
29	38	65.5	393	2	T20268	hypothetical prote
30	38	65.5	697	2	T26707	hypothetical prote
31	38	65.5	787	2	T05617	hypothetical prote
32	37	63.8	44	1	RXRKOT	relaxin - sand tig
33	37	63.8	53	2	C82776	hypothetical prote
34	37	63.8	109	2	PL0233	Ig heavy chain V r
35	37	63.8	112	2	PL0232	Ig heavy chain V r
36	37	63.8	124	2	A21761	high-cysteine chor
37	37	63.8	163	2	T33130	hypothetical prote
38	37	63.8	255	2	T34163	hypothetical prote
39	37	63.8	370	2	JC7998	platelet-derived g
40	37	63.8	370	2	JC7591	spinal cord-derive
41	37	63.8	370	2	JC7592	spinal cord-derive
42	37	63.8	377	2	AD2746	zinc metallopeptid
43	37	63.8	387	2	F82692	conserved hypothet
44	37	63.8	405	2	D81918	sodium-translocati
45	37	63.8	405	2	G81184	Na(+)-translocatin

ALIGNMENTS

RESULT 1

S12205

hypothetical protein 1 (rRNA external transcribed spacer) - mouse

C;Species: Mus musculus (house mouse)

C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 18-Jun-1993

C;Accession: S12205

R;Oberbaeumer, I.

submitted to MIPS, January 1991

A;Reference number: S12205

A;Accession: S12205

A;Molecule type: mRNA

A;Residues: 1-185 <OBE>

A;Cross-references: EMBL:X56974

Query Match 84.5%; Score 49; DB 2; Length 185;

Best Local Similarity 100.0%; Pred. No. 2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGRWCG 8

|||||||

Db 178 GGGRWCG 184

Search completed: April 12, 2005, 18:58:22
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 18:44:33 ; Search time 56 Seconds
(without alignments)
73.154 Million cell updates/sec

Title: PCT-US04-29649A-1
Perfect score: 58
Sequence: 1 CGGGRWCG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	49	84.5	104	2	Q6Z7M0	Q6z7m0 oryza sativ
2	48	82.8	360	2	Q75IX4	Q75ix4 oryza sativ
3	46	79.3	427	2	Q947X5	Q947x5 oryza sativ
4	46	79.3	427	2	Q8H033	Q8h033 oryza sativ
5	45	77.6	273	1	ENV_FLVC1	P21444 feline leuk
6	45	77.6	273	2	Q9J054	Q9j054 feline leuk
7	45	77.6	413	2	Q9LI26	Q9li26 oryza sativ
8	45	77.6	534	1	ENV_FSVST	P03392 feline sarc
9	45	77.6	662	1	ENV_FLVLB	P11261 feline leuk
10	45	77.6	662	1	ENV_FSVGA	P03391 feline sarc
11	45	77.6	662	2	Q7LYY7	Q7lyy7 feline leuk
12	45	77.6	666	2	Q90AE9	Q90ae9 feline leuk

13	45	77.6	668	1	ENV_FLVC6	P21443	feline leuk
14	44	75.9	271	2	Q6K7U5	Q6k7u5	oryza sativ
15	44	75.9	759	2	Q8HZM8	Q8hzm8	equus cabal
16	44	75.9	1968	2	Q8X0C5	Q8x0c5	neurospora
17	43	74.1	112	2	Q6ZAF3	Q6zaf3	oryza sativ
18	43	74.1	118	2	Q43969	Q43969	azorhizobi
19	43	74.1	128	2	Q84PQ1	Q84pq1	oryza sativ
20	43	74.1	129	2	Q6AVV1	Q6avv1	oryza sativ
21	43	74.1	138	2	Q6YYU6	Q6yyu6	oryza sativ
22	43	74.1	151	2	Q6IKQ2	Q6ikq2	drosophila
23	43	74.1	172	2	Q8S1B6	Q8s1b6	oryza sativ
24	43	74.1	262	2	Q9VRS9	Q9vrs9	drosophila
25	43	74.1	266	1	YD74_METJA	Q58769	methanococc
26	43	74.1	274	2	Q84SS0	Q84ss0	oryza sativ
27	43	74.1	305	2	Q8W384	Q8w384	oryza sativ
28	43	74.1	305	2	Q7XD80	Q7xd80	oryza sativ
29	43	74.1	360	2	Q6K612	Q6k612	oryza sativ
30	43	74.1	500	1	DHA1_MOUSE	P24549	mus musculu
31	43	74.1	500	1	DHA1_RAT	P51647	rattus norv
32	43	74.1	500	1	DHAC_RAT	P13601	rattus norv
33	43	74.1	501	2	O35945	O35945	mus musculu
34	43	74.1	507	2	Q80ZX7	Q80zx7	mus musculu
35	43	74.1	528	2	Q6P0N7	Q6p0n7	homo sapien
36	43	74.1	565	2	Q75LV4	Q75lv4	oryza sativ
37	43	74.1	606	1	ABD4_HUMAN	O14678	homo sapien
38	43	74.1	606	2	Q96E75	Q96e75	homo sapien
39	43	74.1	606	2	Q6IAQ0	Q6iaq0	homo sapien
40	42	72.4	33	2	Q6SEW5	Q6sew5	uncultured
41	42	72.4	188	2	Q8S444	Q8s444	sorghum bic
42	42	72.4	332	2	Q9EX11	Q9ex11	streptomyce
43	42	72.4	535	1	ESR1_SALSA	P50242	salmo salar
44	41.5	71.6	222	2	Q65XK5	Q65xk5	oryza sativ
45	41	70.7	56	2	Q656R1	Q656r1	oryza sativ

ALIGNMENTS

RESULT 1

Q6Z7M0

ID Q6Z7M0 PRELIMINARY; PRT; 104 AA.

AC Q6Z7M0;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein OJ1233_A01.24.

GN Name=OJ1233_A01.24;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RA Sasaki T., Matsumoto T., Yamamoto K.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AP004845; BAD15787.1; -.

KW Hypothetical protein.

SQ SEQUENCE 104 AA; 10293 MW; 60BD09F52EE606D6 CRC64;

Query Match 84.5%; Score 49; DB 2; Length 104;

Best Local Similarity 100.0%; Pred. No. 4.1;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGGRWCG 8

|||||||

Db 22 GGGRWCG 28

Search completed: April 12, 2005, 18:58:01

Job time : 59 secs